PaRa: Personalizing Text-to-Image Diffusion via Parameter Rank Reduction

Shangyu Chen & Zizheng Pan & Jianfei Cai & Dinh Phung

MONASH University

Motivation

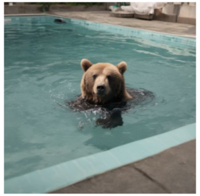
Personalization
≠ New Model ?

What if it's just a subspace of the original model?

"A small bear [V] sitting on a rockery in the park"

"A small bear [V] sitting on the grass"

Input 2 training images of a bear plushie for personalization training.



Generate images with prompts

"A small bear [V] swimming in the swimming pool"

❌ the goal is no longer generative diversity

✅ but **high fidelity to the few training samples**

This can be interpreted as constraining the model's output to a restricted, ***low-dimensional generation subspace***!

Method: Parameter Rank Reduction (PaRa)

Restricting Output Space via Low-Rank Parameter Projection

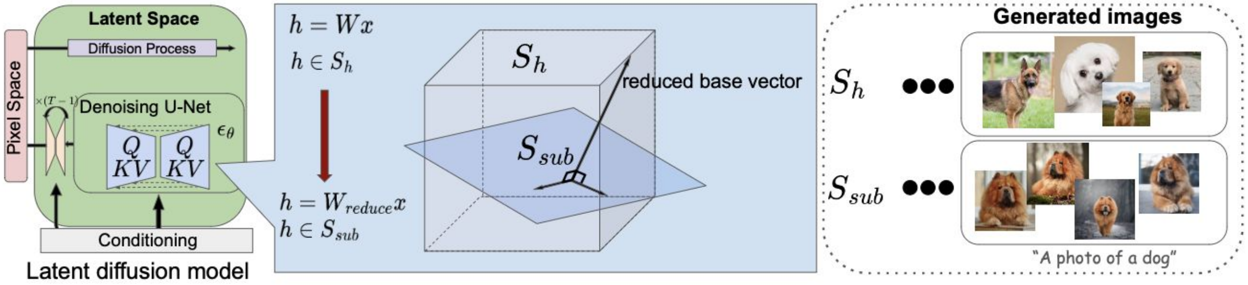
- To align the output with a small set of personalized training images, we propose **PaRa**, a method that **explicitly reduces the model's output space** through **parameter rank reduction**.

- Instead of adding new low-rank components (as in LoRA, where the learned matrices B and A are numerically independent of the pretrained weights), PaRa learns a subspace of the original weight matrix W_0 and projects W_0 onto its orthogonal complement to restrict the output space.

- Given a pretrained weight matrix W_0, we learn a low-rank matrix B and perform:

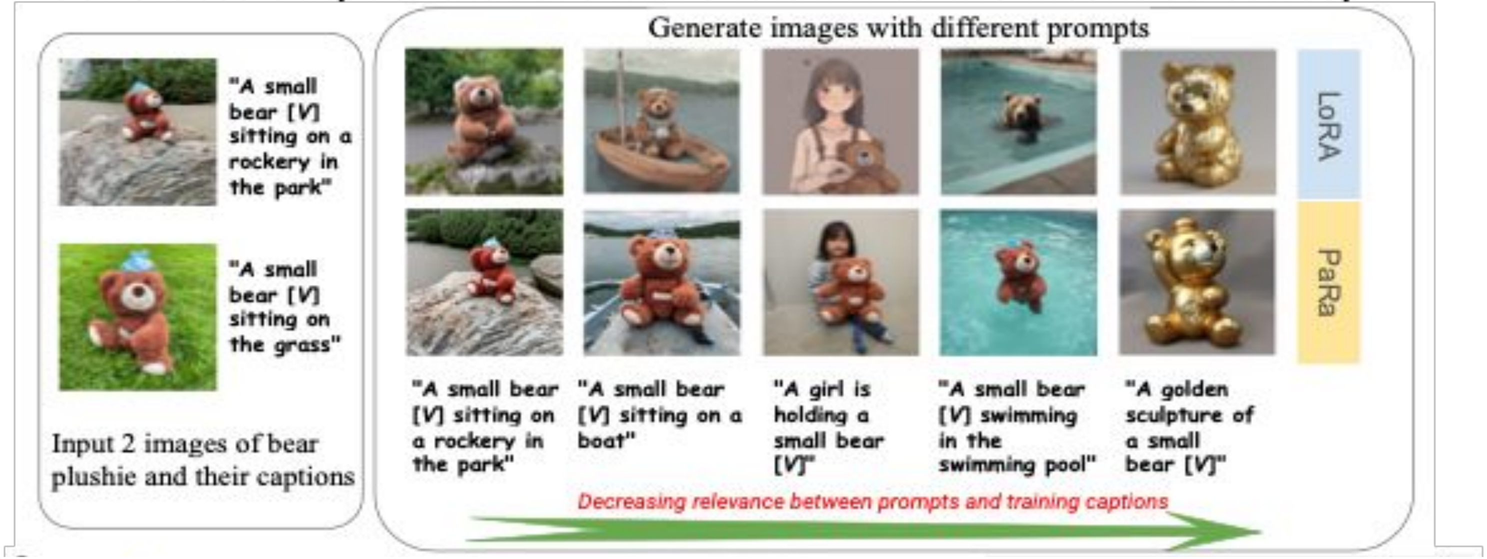
$$W_{\text{PaRa}} = W_0 - QQ^T W_0$$

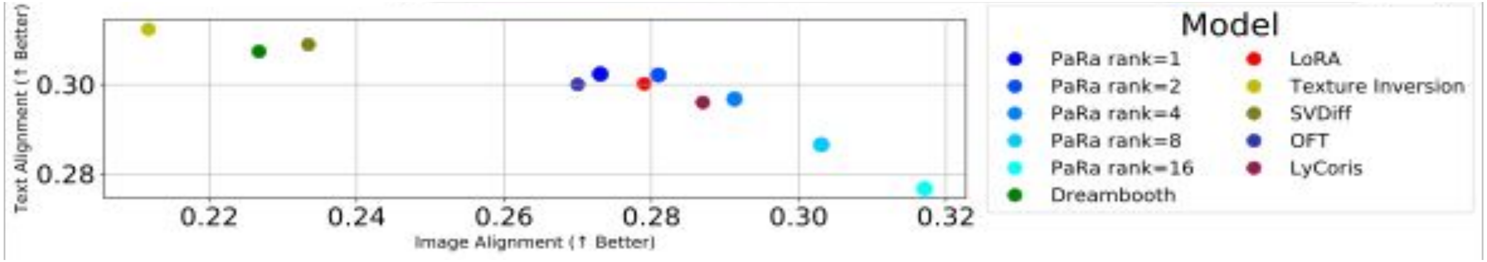
- where $Q, R = \text{QR}(B), \quad B \in \mathbb{R}^{d \times r}, \quad Q \in \mathbb{R}^{d \times r}$ is obtained via QR decomposition
- This removes r dimensions from the output space of W_0.



Key Advantages: Target Alignment

Generates outputs that are more faithful to the intended concept.





Key Advantages: Parameter Efficiency

PaRa requires 2× fewer parameters than LoRA.

	r=2		r=16		r=32		r=128	
Model	PaRa	LoRA	PaRa	LoRA	PaRa	LoRA	PaRa	LoRA
Model Size	1.8 MB	4.8 MB	13 MB	33 MB	22 MB	56 MB	87 MB	190 MB
Train. Time	5.8min	6.2min	6.9min	10.4min	9.4min	15.1min	11.2min	20.1min

Key Advantages: Stable Single-Image Editing

PaRa supports image editing without requiring noise inversion through one-shot learning of the original image and performs generation by directly modifying the prompt.

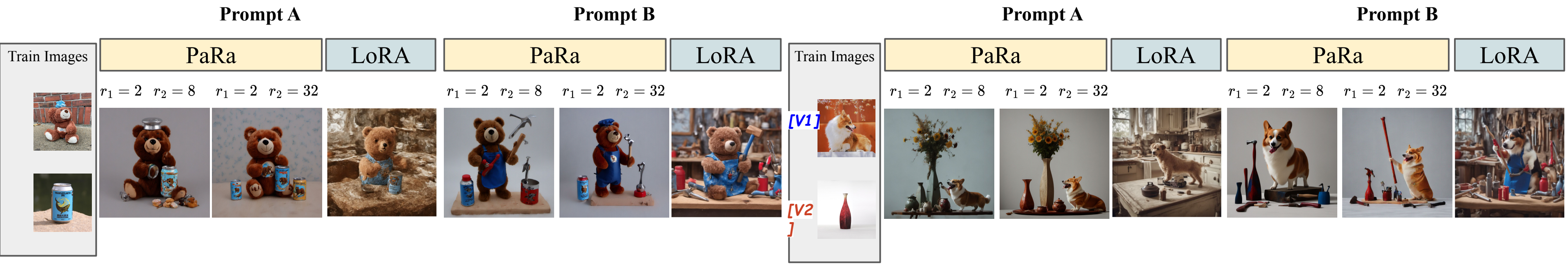
Key Advantages: Multi-Subject Personalization

PaRa supports the combinations of different personalized PaRa models. For two individually trained PaRa models:

$$W_1 = W_0 - Q_1 Q_1^T W_0, W_2 = W_0 - Q_2 Q_2^T W_0,$$

The merged PaRa matrix can be expressed as:

$$W_m = W_1 - Q_2 Q_2^T W_1 = (W_0 - Q_1 Q_1^T W_0) - Q_2 Q_2^T (W_0 - Q_1 Q_1^T W_0).$$



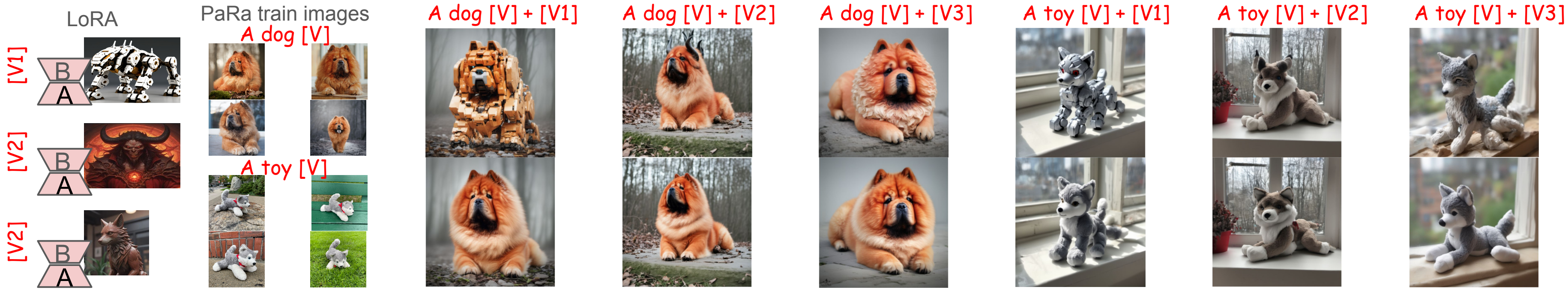
Prompt A: *Concept1 [V1]* is trying to open *Concept2 [V2]* with its paws, located in a nostalgic kitchen filled with vintage furniture and scattered biscuit.

Prompt B: A *Concept1 [V1]* wearing a blue apron is tapping a *Concept2 [V2]* with a miniature hammer standing on a wooden workbench with a red handcraft tool cloth.

Key Advantages: Compatibility with LoRA

For LoRA, the weight is updated as $W = W_0 + \alpha BA$

$\Delta W = (-\alpha_1 Q + \alpha_2 B)(\alpha_1 Q^T W_0 + \alpha_2 A)$, where α_1 and α_2 are parameters controlling the strength of PaRa and LoRA.



Thanks!